Vol. 37 No. 9　　情　報　処　理　　Sep. 1996

解　説

英語と日本語を対象にした文章誤り検出・訂正の共通点と相違点†

Rodney G. Webster††　　中　川　正　樹††

1. はじめに

コンピュータは文字を扱えるようになってから文章作成支援の有力なツールとして大きな役割を果たしてきた．初期においては，文書エディタや文章フォーマッタがこの中心であった．しかし，これらがあたりまえのものとして定着してくると，関心はより高度な支援ツールへと移行してきた．文章校正支援もその1つである．

文章校正支援の研究は大きく2つに分けることができる．誤り検出と誤り訂正である．誤り検出とはコンピュータが文中の「誤り」を見つけ出すことである．どういうものが「誤り」と判定されるかについては後ほど詳しく述べるが，不正な単語，文法の間違いなどが一般的な例である．一方，誤り訂正とは，誤り検出で見つけた誤りに対して，コンピュータが訂正の候補を生成することである．自動的に訂正を行う場合もある．無論後者の方が複雑である．この2つの研究課題は，自然言語処理や人工知能などの技術を応用している．

文章校正支援は，英語から始まった．1960年代にすでに，英語の間違った綴りをコンピュータが自動的に直すことが研究されていた．スペルチェッカが普及している現在，この研究の注目点は単語単位から文単位へと移行した．一方，日本語における文章校正支援に関する研究は，主に1980年代から始まった．これは日本語を扱えるコード体系の実用化，および，文字の入出力技術の開発や普及が遅れたためであった．

本稿では，英語と日本語に対する誤り検出と訂正の研究開発動向について，英語向けと日本語向けを対比しながら紹介する．

2. 誤りの種類について

Kukichは英語における誤り訂正を3つのカテゴリーに分ける[1]．この3つは，“nonword error detection”（単語ではない誤りの検出），“isolated-word error correction”（文脈を利用しない誤り訂正），“context-dependent word correction”（文脈に依存する誤り訂正）である．“nonword error detection”とは，綴りの間違いから存在しない単語が生じた誤りを検出することをいう．これに反して，“word error”とは誤りが混入したにもかかわらず，依然として存在する単語である場合をいう．“nonword error”に対して，70年代初めから80年代初めまで，主にパタンマッチングや文字列比較などが中心に研究された．“isolated-word error correction”とは，文脈を使わずに単語ごとに誤りを検出し，訂正する処理のことである．この研究は1960年代から始まって，今でも続けられている．この研究はミススペル（綴り誤り）の傾向に関する研究とともに発展し，一般的な誤り，または特定の誤りを訂正する様々な手法が開発された．“context-dependent word correction”とは，文全体を対象に誤り検出とその訂正を行うことである．対象となるものには綴りの間違いだけではなく，文法に合わない文，不適切な表現など，様々なレベルのものまで含まれる．これに対する実験は1980年代初めに，自動的な自然言語処理のモデルから始まった．最近は統計的言語モデルが開発されたことで再び注目を集めている．

文章校正支援の立場から，池原らは日本語文に現れる誤りを3つのレベルに分ける[2]．これらは表記レベル，表現レベル，内容/一般常識レベルの3つである．表記レベルの誤りは，たとえば表

† Similarities and Differences between Error Detection and Correction of English and Japanese Text by Rodney G. Webster and Masaki NAKAGAWA (Department of Computer Science, Faculty of Engineering, Tokyo University of Agriculture and Technology).

†† 東京農工大学工学部電子情報工学科

表-1　英語と日本語における誤りの種類

誤りの種類	適用される技術	英文における代表的な例	日本文における代表的な例
表記レベル	形態素解析	ミススペル	表記のゆれ
表現レベル	構文・意味解析	単数型・複数型の誤用	仮名漢字変換誤り
内容/一般常識レベル	意味理解以上	存在しない固有名詞，矛盾する数字，文意など	

記のゆれや俗語/禁止語など，単語そのものが間違っているか不適切なものを指す．表記レベルの誤りに対しては形態素解析技術が利用される．表現レベルの誤りとしては，同音語，類型単語誤りなどが例としてあげられる．これらは，文中の単語は正しい（存在する）ものであるが，その使い方が不適切な誤りである．構文・意味解析技術はこの種類の誤りに対応できる．最後の内容/一般常識レベルの誤りは，単語や文法以外の要素が間違っている文である．例としては，存在しない固有名詞，矛盾する数値や文意などがある．内容レベルの誤りに対応するには，対象知識や専門分野知識を背景とする意味理解以上の技術が必要であることを池原らは報告している．意味理解まで含む自然言語処理はまだ未発達なので，現在のシステムは主に表記レベルと表現レベルを対象にしている．

文章に現れる誤りは，様々な要素から影響を受ける．いくつかの例を次に示す．

（1）　文章を入力した人の熟練度：Grudin はタイプされた英語文に現れる誤りについて，熟練している人が起こす誤りはほとんど文字の挿入であると報告している[3]．一方，素人は文字を入れ替える誤りがほとんどである．

（2）　書かれた言語：言語の特徴によって誤りの種類と傾向は違う．**第3章**では日本語と英語の違いについて詳しく述べる．

（3）　入力デバイス・方法など：（1）で述べたように，タイプされた文章に起こる誤りには傾向がある．打つつもりのキーの周辺にあるキーは間違って打たれる可能性が高い．たとえば，“q”，“w”，“e”，“a”などは“s”を打とうとするときに間違って押される可能性が他のキーより高い．このようなミスタイプ（誤打鍵）は英語においては見落とされがちであるが，日本語を入力する際は仮名漢字変換の段階で確認されることが多い．野田は仮名漢字変換以前の日本語文におけるミスタイプを調査した[4]．野田はこの種の誤りの80%が置換ミス・挿入ミスで，使われていないキーほど誤る可能性が高いと報告している．しかし，これらは入力時に訂正されることが多いので，日本語文ではミスタイプより仮名漢字変換による変換ミス（同音語）の方が問題となる．

表-2　英語と日本語の特徴の対比

英語の特徴	日本語の特徴
文頭に近いほど制約が強い	文末に近いほど制約が強い
分かち書きされている	分かち書きされていない
新語・死語が頻繁に発生	外来語が多い
単語は多義	同音語が多い
語順の制約が厳しい	語順の制約が緩やか
単数型・複数型	相当するものがない
動詞と名詞の交換性	交換性がない
前置詞	助詞
英国式/米国式のスペル	1つの単語には表記が多数
	語彙が大きい
	複合名詞が多い

3.　英語と日本語における自然言語解析の特徴

英語は長い間いくつもの外国語から言葉や文法などを取り入れながら発展してきた言語である．そのため，文法や綴りなどが統一されていない点が多く，外国語として勉強している人を悩ませる．英語の特徴は**表-2**に示す．

文頭に近いほど制約されていることは，文の解析は文頭から始めた方が有効であることを示す．また，分かち書きされているということは，単語抽出が容易であり，辞書を用いて未登録語を検出することが有効に利用できることを示す．しかし，この分かち書きされていることにも特有の誤りが生じることがある．英語文では空白の欠落や挿入など，単語の境目が誤りとなることも少なくない．Kukich の研究によると，40,000 語の文章には単語とならなかった綴り誤りのうち，15%がこのような誤りであった[5]．空白の欠落はたとえ

ば，“of the”が“ofthe”となるような誤りである．空白の挿入はたとえば，“forgot”が“for got”になるような誤りである．この例で見られるように，誤りが起きても存在する単語になってしまう誤り，つまり“word error”が問題になる．Mittonの研究によると，空白の欠落や挿入などの誤りから1つ以上の存在する単語ができることは多い[6]．これが原因でその誤りが見つからないことも問題である．タイプ入力では空白が欠落しやすく，OCRでは誤りの空白が挿入しやすい[5]．

単数形・複数形は辞書を用いた未登録語の検出を惑わす要素である．単数形・複数形の場合，ほとんどの言葉は簡単な規則で単数形から複数形が推定できる．しかし，何の規則にも従わない例外もある．たとえば，“die”（サイコロ）→“dice”，“mouse”→“mice”などがある．

名詞と動詞の交換性は品詞の推定や意味の推定，未登録語に対処するための処理を妨げるものである．また，構文解析にも影響を与えて正しい文が誤りと判定される原因となる可能性がある．この問題を解決するには特別な配慮が必要である．

日本語は，近代において外来語が多く導入されていることを除けば，比較的独自に発展してきた言語である．文法的な面ではほとんど規則に沿い，動詞の活用には例外がわずかしかない．日本語の特徴は表-2に示す．

日本語は文末に近いほど制約されているため，英語とは反対に文末から抽出した方がいい．野美山らは日本語文を解析することについて，文末に近いほど，係受けの候補数が少なくなると書いている[7]．また，「は」「を」「へ」などの助詞は語間の区切りを示すので，その前方にある言葉の抽出に利用できるうえ，品詞の推定などにも役立つ．助詞の不適切な使用も，誤り検出の対象となる．菅沼らの研究は副助詞の「は」に注目した[8]．日本語文では，一文中に助詞の「は」，もしくは「は」という音（wa）が複数個含まれることが多い．しかし，このような文は読みにくいので，なるべく避けた方がよい．彼らは副助詞の「は」を効率良く検出する方法を発表している．

飯盛らの研究によると，日本語は外来語が多く，そのほとんどが名詞成分であり，約8割は英語から来ている[9]．外来語は辞書に登録されていない地名や人名などが多いが，語尾に「駅」や「先生」などが付いていることから判定できることもある．日本語における外来語で特に問題となることは，一語に対して表記法（異なり語）が複数存在することであると島津らは指摘する[10]．たとえば，リスポンス→レスポンス(発音の違い)，コード→コーディング（動詞の活用）などが問題となる．島津らはこの問題を解決するために異なり語の自動生成を研究している．異なり語の自動生成は入力片仮名文字列から異形表記を作成する．この異形表記は誤り訂正に利用される．

専門の文章に出てくる学術用語なども問題となる．辞書が対象となる分野に適応していないと，これらの言葉は無条件で誤りとなる．納富らは日本語において次の2つの点に注目して研究している[11]．（1）専門用語(未知語)の多くは複合名詞であること，（2）複合名詞は構造を持つこと，である．複合名詞は（修飾—被修飾）という形で，（形容詞—名詞），（副詞—動詞）などの係受け関係を表現したものが多い．これは，表現を簡略化する省略形式の1つであるという見方もある．一般に省略の対象となるものは付属語要素，助詞，用言の活用語尾などである．納富らはあらかじめ名詞や動詞などの付属語要素を想定し，その情報を使って複合化された単語からできる文が成り立つかどうかを調べた．

日本語は漢文から取り入れた単語が多いが，日本語の母音/子音数は中国語に比べると少なく同じ発音になる単語が多い．したがって同音語が多くて，仮名漢字変換の変換ミスにつながることが多い．たとえば，「いし」という発音を持つ単語は「意志」「石」「医師」などたくさんある．さらに，1つの動詞でも微妙なニュアンスを伝えるために，それに当てられる漢字が多数あることもある．たとえば，「きる」という単語には「切る」「斬る」「伐る」がある．仮名漢字変換を用いた入力で，変化ミスが起こる可能性は高い．また，仮名漢字変換による入力では区切りの位置を間違えることもある．

日本語では1つの単語に複数の表記があることも問題となる．日本語では漢字，平仮名，片仮名，英数字などと様々な文字が使われる．また，「行う」と「行なう」のように送り仮名が2つ以上存在する単語もある．特に問題となるのは表記の

「ゆれ」，すなわち，1つの文章中で単語の書き方が統一されていないことである．さらに，この「ゆれ」は文章の書き方，句読点の使い方などでも問題となる．「ゆれ」の一般的な例としては，同じ文章で「する」と「します」を書くこと，「．」と「。」の句読点を両方使うことなど様々なものがある．これらは，全体的に文章のスタイルを統一する目的で，誤り検出の対象となる．日本語のゆれの問題は，英語におけるスペルの統一に相当する．たとえば，“colour”（英国式）か“color”（米国式）に統一するなどである．

日本語は分かち書きされていないので単語抽出が比較的難しい．しかし，単語抽出の処理を手助けするいくつかの要素はある．たとえば，上記で述べた日本語の特徴である助詞，漢字と仮名の境目，句読点などを利用する処理で単語の区切りが想定できる．抽出された単語は辞書や単語表と照合することによって確認できる．漢字仮名混じりの文では，形態素解析より精度良く単語の分割処理を行うことができる．荒木らは，べた書き仮名文の場合にはマルコフ連鎖モデルの使用が有効であることを指摘している[12]．

英語では主語，動詞，そして，目的語のように，語順はある程度限られている．しかし，日本語では文法が語順に与える制約は緩やかである．基本的には動詞が最後にあることと，修飾語の後に被修飾語があること以外の制約はない．これは構文解析にとって望ましくない特徴である．語順の制約が大きいと，その制約が単語の意味や品詞などの推定に役立つ．そうでない場合には，前述の単語抽出のように助詞を利用することや辞書との照合で対応するしか方法はない．

4. 英語と日本語の文章校正支援

4.1 誤り検出の研究

誤り検出で，最も多く研究されている手法は形態素解析を利用する方法である．形態素解析には辞書を用いる．文中から抽出された単語と辞書を照合して，なければその単語は誤りとされる．辞書照合の分野では，検索方法，パタンマッチングや文字列の照合などが研究の中心である．文章に現れる誤りに十分に対応するには，それなりの大きさの辞書が必要となる．しかし，辞書があまり大きくなると検索に時間がかかるうえ，メモリも必要になる．

辞書を用いる処理では未登録語が大きな問題となる．この問題を解決するため，山田らは英文における未登録語の意味推定を研究している[13]．また，山田らは，英語文は分かち書きされているためその抽出と構文解析が日本語文より容易に処理できることに対して，英語の単語は多義であることが多く意味解析において処理がきわめて難しくなることを指摘している．

n-gram モデル（文字遷移確率モデル）は誤りを検出するために用いられる手法である．n-gram とは n 個の文字や言葉からなる列である．このような文字列や言葉の列に対して，出現頻度などの統計的なデータを表としてあらかじめ用意する．対象となる文の中の各 n-gram は表と照合され，存在しない n-gram や出現回数が少ない n-gram は誤りである可能性が高いとして指摘される．n は一般に 1～3 である．英語文に対して n-gram は文字数が少ないためかなり有効な手法である．日本語には，単語の切れ目が明瞭ではない，文法が語順に与える制約が緩やかである，一般に語彙が大きい，という3つの特徴がある．これらから，日本語文に対して n-gram は英語ほど有効ではないことが予測されている．しかし，予想に反して，丸山らは英語と同程度の効果を報告している[14]．丸山らは，十分な効果が得られるためには大量のデータから統計的な情報を集める必要があることも指摘している．

漢字に対して n-gram は相当量のメモリを要求する．そのうえ，仮名に対しては n を大きくとらないと十分な制約にならないということも指摘されている．これに対して，伊東は n-gram に品詞を属性として導入し，仮名に対してはより強い制約を与えること，漢字は品詞ごとのマクロ文字にグルーピングし縮退させることを提案した[15]．

形態素解析では文節以上のものは扱わないが，誤り指摘規則で文節間の関係を処理することができる．願化らはこの誤り指摘規則について報告している[16]．この研究では誤り指摘の知識を大量のデータから抽出し，コンピュータが扱える誤り指摘規則に変換する．誤り指摘規則を利用する処理では，抽出された知識から生成された規則が，実際のデータにうまく使えるかどうかが問題とな

る．願化らは，一般規則では文法的な間違いは検出できるが，「話し会う」のような文法としては正しい間違いを逃してしまうと報告している．このような誤りを拾うには，語彙規則が必要となる．一般規則は文法に対するチェックを行うものであり，語彙規則は単語個別に存在する規則をチェックするものである．

形態素解析以前の基本的手法として字面解析がある．これは辞書照合や文法解析などを用いない．応答時間を重視した場合，字面解析が利用できる．日本語では受身形の動詞は受身，可能，自発，尊敬という意味が持てるため，誤りとなる可能性が高い．牛島らは字面解析を用いて受身形の使い方を指摘するツールを開発した[17]．

文節以上の要素を対象にした誤り検出には構文解析が必要となる．前述したように，英語文ではミススペルによって存在する単語が生じる可能性が高い．Mittonは学生の手書きのレポートを925個調べ，ミススペルの40%もが存在する単語になっていたことを報告した[6]．また，Wingらは1185個のミススペルのうち約30%が，存在する単語になっていると報告した[18]．これは英語文における構文解析の重要性を示している．代表的な手法では，主語と動詞の活用形が合っていないことや語順がおかしいことなどを調べる．

4.2 誤り訂正の研究

誤り訂正では，対象となる誤りが習慣性のものかそうでないかによってその難しさが違う．習慣性のある誤りに対して，訂正の候補は誤り検出の結果/処理に基づいて推定されることが一般的である．

英語文における表記レベルの誤り，つまり単語のミススペルに対して行われる誤り訂正の手法で，最も単純なものは辞書処理である．習慣性のないミススペルに対して利用できる手法は，入力単語と辞書にある単語の距離を計り，近いものを候補として選ぶ．しかし，これは入力の単語と辞書にあるすべての単語との距離を計算する手間が大きく，あまり実用的ではない．Takahashiらはこの点を指摘し，OCRの認識に対する後処理における処理速度改善のために，誤り訂正の候補を生成する手続きを2つの段階に分ける手法を提案した[19]．この手法ではまず辞書から訂正の候補が簡単な判定で抽出される．次に入力の単語と候補の距離が計られる．

習慣性のある誤りは誤り指摘規則で対応することが一般的である．誤り指摘規則で発見された誤りは規則に基づいて訂正の候補を生成することができる．これは表記レベル，表現レベル，内容/一般常識レベルでも起こる誤りに共通する特徴である．

表記レベルの誤りとして，英語で書かれた文章に起こるミススペルに対して多くの調査が行われてきた．収集された情報は，誤りから訂正候補を生成するために役立つ．一般的に訂正の手法には，音韻学の知識を用いて誤りと候補の距離を計るか，文字の挿入や置換などに重みを付けて誤りから候補を生成してその距離を計るものがある．また，よく見られるスペルの間違いに基づいたヒューリスティックなものもある．日本語の場合も同様で，同音語，ふり仮名などは誤り指摘規則によって検出できるため訂正の候補が生成しやすい．しかし，こういった手法は習慣性のない表記レベルの誤りに対して利用できないという欠点を持つ．

日本語文における誤り（誤字，脱落，誤挿入）に対して，2重マルコフ連鎖モデルを用いて誤りを検出・訂正する方法はその有効性が知られている．荒木らはさらに3重マルコフモデルの利用を研究している[20]．この研究は，正しい文より誤りの文はマルコフ連鎖確率の値が小さいと仮定する．マルコフ連鎖確率の値から誤りの場所と種類が求められる．誤り訂正も同じマルコフモデルを用いて行われる．たとえば，誤りが脱落であれば候補の文字が挿入され，再計算されたマルコフ連鎖確率の値から候補が正しいかを計る．荒木らはマルコフモデルを用いた手法を既存のスペルチェッカと比較する実験を行った．後者は誤り検出に優れていることに対して，誤り訂正にはマルコフモデルを用いた方が有効であると報告している．

5. おわりに

内容/一般常識レベルで起こる誤りに関しても，単純な誤りなら上記の技術でも十分対応できる．たとえば，存在しない固有名詞，矛盾する数字などには形態素解析のような低レベルの手法がよく用いられる．しかし，文意の問題などは，自然言語処理の意味理解などが必要となる．

誤り検出・訂正の現状を見ると，様々な手法や技術が研究されているが，研究の課題はまだ多く残されている．今後の進展のためには，これまで研究開発された技術を公正に評価することがまず第一に必要である．そのためにはより大量のコーパスが不可欠である．日本語の研究において言語解析や手法の評価を行うために必要な一般公開された文章は多くない．

ツールのユーザインタフェースを改善する必要性も高い．下村ら[21]は人間の誤り検出能力と，機械による誤り検出の人間にとっての効果について実験を行い，機械がすべての誤りを検出できなくても，ある種の誤りを特に指摘すべきであると述べている．

英語では30年間も，日本語でも10年間以上も研究されてきた文章校正支援はいまだにその普及が進んでいない．新聞社や出版社など，企業向けのシステムはすでに開発されて実用化されているが，一般的なユーザへの普及はそれほどではない．今後は，既存の技術の評価とユーザインタフェースの研究が重要になろう．

謝辞 本文の内容について有益な助言をいただいた西村恕彦教授に感謝する．

参 考 文 献

1) Kukich, K.: Techniques for Automatically Correcting Words in Text, ACM Computing Surveys, Vol. 24, No. 4, pp. 377-439 (1992).
2) 池原 悟，小原 永，高木伸一郎：文書校正支援システムにおける自然言語処理，情報処理，Vol. 34, No. 10, pp. 1249-1257 (Oct. 1993).
3) Grudin, J.: Error Patterns in Skilled and Novice Transcription Typing, in Cognitive Aspects of Skilled Typewriting, W. E. Copper, Ed. Springer Verlag, New York (1983).
4) 野田雄三：誤打鍵特性の調査と分析，情報処理学会第47回全国大会論文集，3，1W-3，pp. 217-218 (1993).
5) Kukich, K.: Spelling Correction for the Telecommunications Network for the Deaf, Comm. ACM, Vol. 35, No. 5, pp. 80 90 (1992).
6) Mitton, R.: Spelling Checkers, Spelling Correctors, and the Misspellings of Poor Spellers, Information Processing Management 23, 5, pp. 495-505 (1987).
7) 野美山浩：大規模日本語テキストからの依存構造の抽出，情報処理学会第49回全国大会論文集，3，7G-2，pp. 179-180 (1994).
8) 菅沼 明，松尾 朗，牛島和夫：日本語文章推敲支援ツール『推敲』における字面解析 助詞「は」に着目して―，情報処理学会第39回全国大会論文集，3J-5，pp. 781-782 (1989).
9) 飯盛可織，佐川雄二，大西 昇，杉江 昇：大規模コーパスに基づいた未登録語の傾向分析，情報処理学会第49回全国大会論文集，3，7G-7，pp. 189-190 (1994).
10) 島津美和子，吉村裕美子，平川秀樹，天野真家：片仮名異形表記・誤記修正機能の開発・評価，情報処理学会第44回全国大会論文集，3，3Q-4，pp. 249-250 (1992).
11) 納富一宏，中条和光，石井博章，内山明彦：日本語文書校正支援ツールの開発―複合名詞の統語的検定について―，情報処理学会第49回全国大会論文集，3，3S-7，pp. 291-292 (1994).
12) 荒木哲朗，池原 悟，土橋潤也：べた書きかな文の仮文節境界の補正方法，自然言語処理，98-1，pp. 1-7 (1993).
13) 山田一郎，山村 毅，佐川雄二，大西 昇，杉江 昇：英文における未登録語の意味推定の検討，自然言語処理，93-9，pp. 63-70 (1993).
14) 丸山 宏：Nグラムモデルによる，日本語単語の並び替え実験，情報処理学会第49回全国大会論文集，3，7G-3，pp. 181-182 (1994).
15) 伊東伸康：Bigramによるオンライン漢字認識の文脈後処理手法，自然言語処理，97-6，pp. 37-44 (1993).
16) 願化真志，小山紀子，斎藤宏美：文章作成支援システムにおける日本語処理(2)―文章誤り指摘規則―，情報処理学会第35回全国大会論文集，4S-6，pp. 1263-1264 (1987).
17) 牛島和夫，石田真美，Jeehee Yoon，高木利久：日本語文章推敲支援ツールにおける受身形の抽出法，情報処理学会論文誌ショートノート，Vol. 28, No. 8, pp. 894-897 (Aug. 1987).
18) Wing, A. M. and Bradley, A. D.: Spelling Errors in Handwriting: A Corpus and Distributional Analysis, in Cognitive Processes in Spelling, U. Frith, Ed. Academic press, London (1980).
19) Takahashi, H., Itoh, N., Amano, T. and Yamashita, A.: A Spelling Correction Method and its Application to an OCR System, Pattern Recognition, Vol. 23, No. 363-377, pp. 363-377 (1990).
20) 荒木哲朗，池原 悟，塚原信幸，小松康則：マルコフ連鎖モデルによる仮名文と英語文の誤り訂正，情報処理学会第48回全国大会論文集，3，1Q-8，pp. 51-52 (1994).
21) 下村秀樹，並木美太郎，中川正樹，高橋延匡：人間の文章誤り検出能力と検出機能の効果に関する実験，情報処理学会論文誌，Vol. 33, No. 12, pp. 1607-1617 (Dec. 1992).

(平成7年9月29日受付)

Rodney G. Webster（正会員）

1969年生．1990年タスマニア大学人文学部卒業．1994年東京農工大学大学院工学研究科電子情報工学専攻博士前期課程修了．同年同大学博士後期課程に進学，1997年修了予定．現在，手書き文字認識および文脈処理の研究に従事．

中川　正樹（正会員）

1977年東京大学理学部物理学科卒業．1979年同大学院修士課程修了．同在学中，英国Essex大学留学（M. Sc. in Computer Studies）．1979年東京農工大学工学部助手．現在，助教授．日本語計算機環境，日本語文書処理，手書きインタフェースなどの研究に従事．理学博士．